Helping You Understand Localization

How translation memory really works

INTRODUCTION

This paper is intended for program managers, project managers, engineering leads, documentation writers or anyone, who will be involved in a localization project. By use of a hypothetical example, the paper describes the process of localization. It breaks down the process into phases and explains who, what, where and how is involved in a successful localization project.

KEYWORDS: *translation, localization, internationalization, simultaneous release, deployment strategy, request for proposal, localization costs, translation leverage, translation memory, localization schedule, translation drops, online collaboration portal, EzBIS*

First: What does L10N Stand for?

It is an industry recognized abbreviation for the term localization. There are 10 letters between the letters L and N.

Kitzland and its localization needs

Here is a hypothetical scenario. Let's suppose that a US based company, let's name it *Kitzland*, is developing an educational web-based application for children. It will be an interactive website, through which children can play games, learn songs, and access information on interesting topics such as science, the animal kingdom, travel etc… *Kitzland* plans to make this site available in a variety of languages: US English, French, Italian, German, Spanish, Brazilian Portuguese, Japanese, Korean, Chinese and Hebrew.

One of the songs that *Kitzland* will teach children worldwide is the popular kids' song "Old MacDonald had a farm".

Old Macdonald had a farm, E-I-E-I-O
And on his farm he had a cow, E-I-E-I-O
With a "moo-moo" here and a "moo-moo" there
Here a "moo" there a "moo" Everywhere a "moo-moo"
Old Macdonald had a farm, E-I-E-I-O

Old Macdonald had a farm, E-I-E-I-O
And on his farm he had a pig, E-I-E-I-O
With a (snort) here and a (snort) there
Here a (snort) there a (snort) Everywhere a (snort-snort)
With a "moo-moo" here and a "moo-moo" there
Here a "moo" there a "moo"

Everywhere a "moo-moo"
Old Macdonald had a farm, E-I-E-I-O

Old Macdonald had a farm, E-I-E-I-O
And on his farm he had a horse, E-I-E-I-O
With a "neigh, neigh" here and a "neigh, neigh" there
Here a "neigh" there a "neigh" Everywhere a "neigh-

neigh"
With a (snort) here and a (snort) there Here a (snort)
there a (snort)
Everywhere a (snort-snort)
With a "moo-moo" here and a "moo-moo" there
Here a "moo" there a "moo" Everywhere a "moo-moo"
Old Macdonald had a farm, E-I-E-I-O

The text of this song appears in a separate HTML file titled "oldmacdonald.htm" and this file is a part of the delivery to EzGlobe. We will use this file as an example throughout this paper to illustrate the process of localization.

Internationalization vs. localization

Before we get started, let us first explain two important terms that people often use interchangeably which is incorrect: internationalization and localization.

Internationalization, often referred to as I18N, is an integral part of the product design. It is the process of making a product flexible so that it can successfully function on non-native operating systems and handle multiple languages and cultural conventions. It is a prerequisite for successful localization.

Localization, often referred to as L10N, is the process of cultural and linguistic adaptation of products and services for foreign markets.

In this paper, we will mostly talk about localization but we may refer to internationalization from time to time.

When should Kitzland start thinking about localization?

The simple answer is the earlier the better. Localization is an intricate process and, depending on its complexity and scope, it can last several months. Therefore, *Kitzland* must decide during the product planning cycle if they want to release the localized variants simultaneously with the English or if they want to first release the English version

and follow with the foreign variants afterwards. Depending on their decision, they would need to start localizing at different times during their development cycle. The following is a graphical illustration of *Kitzland*'s choices.

Scenario A: Simultaneous deployment

Scenario B: Staggered deployment
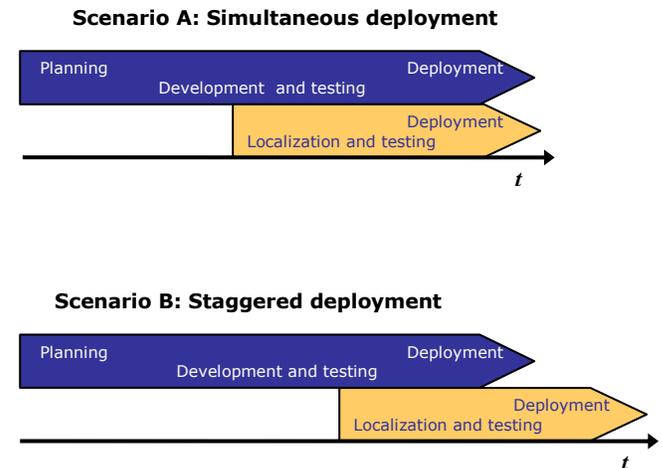
Kitzland decides to release all languages simultaneously.

> **Deciding on the deployment strategy early allows for appropriate timing of the localization activities.**

Requests for Proposal (RFPs)

Kitzland does not have any localization resources in-house, so they will have to rely on outside vendors. After their primary research, they select a handful of professional localization companies and request them to bid on the project. The best way to have an accurate idea of the cost is to provide actual files to the localization companies for detailed analysis.

For those who are new to localization, the proposals may be confusing. They will likely contain a brief description of the proposed localization process, a proposed high-level schedule and the quote. The quote will have multiple of line items. Charges for 100%

matches (exact matches), fuzzy matches, engineering, glossary creation and maintenance, translation memory related charges, and project management fee are just a few examples. The complexity of the selection process, however, does not stem from the simple comparison of the individual line items. Depending on the methodology and computer-aided translation (CAT) solutions the vendors use, each vendor may come up with different analysis results. It is therefore very important not to look just at the price per word, for example, but to closely evaluate the result of the analysis. The better translation leverage you can obtain, the lower will be your translation cost and greater the quality and consistency. Let us explain by using our first concrete example.

As we mentioned earlier, one of the Kitzland's files is "oldmacdonald.htm". The following are results of word count analysis performed by using two different translation memory[1] (TM) solutions:

Translation memory line item	TM System A	TM System B
Repetitions	108	108
In-context matches	0	0
100% matches	0	0
95% - 99% matches	0	0
85% - 94% matches	0	26
75% - 84% matches	0	12
50% - 74% matches	0	20
No match	107	49
Total words	215	215

Even though both systems returned exactly the same number of total words, there are other differences and TM system B appears to be more efficient. As you can see in the Old MacDonald song, there are 50% repetitions (repeated sentences), but also

---

[1] A Translation Memory is a database that stores translated segments for future reuse. It also recognizes repetitive and similar segments within text and these segments can be reused as well.

what we call "fuzzy matches" (approaching sentences).

For example, the sentence "And on his farm he had a cow, E-I-E-I-O " is a fuzzy match of "And on his farm he had a pig, E-I-E-I-O". Fuzzy match prices are discounted based on their percentage of similarity to the sentence they resemble. The difference between the above 2 sentences is 14%; it is therefore an 86% match on which a 40% discount[2] is applied to the full word price.

In other words, without taking any of this into account, Kitzland would pay for 215 words. Now, with TM solution B, they would pay for only 123.6 words. That is a 42.5% saving in this case!

Even better, the translation memory database is built over time so the more documents *Kitzland* translates, the more savings they realize.

Although important, the savings are not the only reason why using an efficient translation memory solution is valuable. The consistency of the translation is another important aspect.

> **Choosing the right translation memory solution means higher translation reuse leading to a decreased cost and greater consistency.**

Kitzland finally selects EzGlobe as their localization partner.

EzBIS

Now that *Kitzland* has become EzGlobe's client, they will gain access to EzBIS -- EzGlobe's proprietary online collaboration portal. EzBIS is the main artery of the project as it connects of the players and

---

[2] EzGlobe applies 40% discount but other vendors may have different pricing policies.

allows them to collaborate efficently. From now on, all project-related information and documents will flow through this system.

Three phase process

Although not one localization project is alike, all of them typically have three main phases:

1. Project setup and planning
2. Project execution
3. Project evaluation and improvements implementation

In the following sections, we will explain what happens in each of these phases.

PROJECT SETUP AND PLANNING

First, EzGlobe needs to fully understand the client's needs. A team of EzGlobe specialists, such as project manager, linguist, and engineer work with *Kitzland* to design a detailed project plan. This project plan outlines how many times, when and by which source controlled method *Kitzland* will deliver their files to EzGlobe for localization processing, so that the ultimate goal of deploying all languages simultaneously is not jeopardized.

Further, the plan will identify the necessary resources both at *Kitzland* and EzGlobe.

Finally, given the project's scope, available resources and *Kitzland*'s needs, EzGlobe will create a detailed project schedule.

PROJECT EXECUTION – DROP 1

Kitzland's team has developed approximately 90% of the content and they are ready to deliver their files to EzGlobe.

EzGlobe's dedicated localization engineer runs the files through a preproduction cycle. He will analyze and convert them to a translation memory friendly format. Let us again illustrate on the oldmacdonald.htm example.

The following is an excerpt from the oldmacdonald.htm code.

```html
<html>
<head>
<meta http-equiv="Content-Type"
content="text/html; charset=windows-
1252">
<title>Old MacDonald had a
farm</title>
<link rel="stylesheet" type="text/css"
href="../../Css/Kitzland.css">
</head>

<body>
<pOLD MACDONALD HAD A FARM<hr></p>
<p class="Orange">Old Macdonald had a
farm, E-I-E-I-O <br>
And on his farm he had a cow, E-I-E-I-
O <br>
With a "moo-moo" here and a "moo-moo"
there <br>
Here a "moo" there a "moo" Everywhere
a "moo-moo" <br>
Old Macdonald had a farm, E-I-E-I-O
</p>
```

As you may imagine, leaving the html tags as-is would lend itself to extreme inefficiencies and would create a high risk of file corruption. So, EzGlobe's engineer uses a toolset to segment the text in sentences, highlight the text to translate, and to fade and protect the html tags against inadvertent alteration or deletion.

While the engineer is preprocessing the files a dedicated language lead creates a glossary of the most frequently used terms in the file set *Kitzland* delivered. EzGlobe translates the glossary first and uploads it to EzBIS so it is available to all stakeholders (not only the translators) 24/7. The advantage of using the glossary is to assure that all translators on a particular job translate the terms consistently throughout the project.

Finally, EzGlobe's project manager uploads all of the components of the so called translation kit to EzBIS and gives the green light to the translators to start.

Naturally, Oldmacdonald.htm is not the only file that needs to be translated. As a matter

of fact, the total volume to translate is 90,000 words. Considering *Kitzland*'s schedule, EzGlobe will deploy three translators for each language on this job. This is 27 translators in total! Luckily, thanks to EzGlobe's translation memory solution these translators can reuse each other's work to propel the work to the successful end. Even more importantly, this solution allows them to keep the translation style and terminology consistent.

During the translation, EzGlobe's project manager will keep tight control over the work via EzBIS.

Once the files have been translated, it is EzGlobe's turn to perform quality control check. There are several quality control methods and *Kitzland* has selected the default method of cross translator sample check.[3] EzGlobe then delivers the translated and QAed files to *Kitzland. Kitzland* now has the option to provide their feedback on the translation. EzGlobe will implement it free of charge.

PROJECT EXECUTION – DROP 2

By now, *Kitzland* will have finished their content development and *Kitzland* gives the green light to EzGlobe to finish the translation.

The previously described preproduction process will repeat itself; however, the difference is that a multilingual translation memory already exists. The translators will translate only the new or modified sentences.

Although the number of file drops is unlimited, it is best to keep them to a minimum as each drop is associated with extra cost. However, sometimes, adding an extra drop may help shorten the overall schedule and if early deployment is

desirable, the extra drop may be worth the investment. It is therefore important to find a fine balance between the number of drops and schedule constraints.

> **Keeping the number of localization drops to a minimum decreases the overall cost of a localization project.**

PROJECT EXECUTION – POSTPRODUCTION

This is the final stage of production.

At this point, EzGlobe's team will convert the files back to their original format and will compile them for the QA phase.

For documentation files (e.g. Adobe FrameMaker or Microsoft Word) EzGlobe will perform desktop publishing (DTP). If desired, EzGlobe will create post script files or PDFs ready for publishing (printing or publishing online).

Finally, *Kitzland* will have the opportunity to review the files for accuracy and validate them before EzGlobe's final handoff.

PROJECT EVALUATION AND FURTHER PLANNING

In most cases, no matter how thorough the planning, there will be challenges that the team will have to address during the life of a project. This is largely due to the fact that most of the projects are unique and unique situations will confront the team.

Stopping the project at the final handoff would be a mistake. There are always valuable lessons to be learned and unless the team takes the time to analyze them and propose executable solutions expensive errors will come back.

---

[3] Other methods of linguistic quality control are: cross translator full review and review by a specialist. To learn more about these methods, please contact us at solutions@ezglobe.com.

EzGlobe acts as a strategic partner for companies that believe in the importance of addressing their clients, partners or employees in their own language. The company helps its clients go global by providing **professional translation, localization and internationalization services**.

www.ezglobe.com

Sophia Antipolis, France	**Massachusetts, USA**
Tel.: +33 4 92 94 23 90	Tel.: +1 781 322 0370